

# IFCS-2004 Program

## Wednesday, July 14

### 9:00-12:00 Short Courses

Shortcourse 1: PLS Latent Variables Soft Modelling

Instructor: Vincenzo Esposito Vinzi, Dipartimento di Matematica e Statistica, Università degli Studi di Napoli “Federico II” **Location: Alumni Lounge, HUB**

Shortcourse 2: Introduction to Microarrays and Clustering

Instructor: Bill Shannon, Washington University in St. Louis School of Medicine **Location: Hermann Lounge, HUB**

### 1:00-4:00 Short Courses

Shortcourse 3: Axiomatics in Bioconsensus

Instructor: William H. E. Day, Port Maitland, Nova Scotia, Canada  
**Location: Alumni Lounge, HUB**

Shortcourse 4: Latent Class Models for Clustering and Classification

Instructors: Jay Magidson, Statistical Innovations, Belmont MA, and Tony Babinec, AB Analytics, Chicago IL. **Location: Hermann Lounge, HUB**

# Thursday, July 15

**8:45-9:10** Opening Remarks **Ballroom, HUB**

**9:10-10:10 a.m.** Plenary **Ballroom, HUB**

Chair: Mel Janowitz, Rutgers University

Jim Ramsay, McGill University, “Modeling the Evolution of Disease Symptoms with Principal Differential Analysis”

**10:30-12:30** Sessions

I-1: Data Analysis Issues in PLS Regression and Path Modeling, **MTCC, Ballroom**

Chair/Org.: V. Esposito Vinzi, University of Naples “Federico II”

Michel Tenenhaus, HEC Business School, “PLS Path Modeling for Multiple Table Analysis”

Gilbert Saporta, CNAM Paris, and Cristian Preda, Université de Lille, “Clusterwise PLS Regression with Applications to Functional Data”

Xinli Bao and Hamparsum Bozdogan, University of Tennessee at Knoxville, “Subsetting Kernel Regression Models Using Genetic Algorithm and the Information Measure of Complexity”

I-2: Graphics in Statistics, **Ballroom, HUB**

Org.: Leland Wilkinson, SPSS, Inc.

Chair: Mike Larsen, Iowa State University

Di Cook, Iowa State University, “Visualizing Class and Cluster Structure beyond 3D”

Graham Wills, SPSS, Inc., “Transformations of Graphs”

Leland Wilkinson, SPSS, Inc., “Streaming Data in Multiple Windows”

I-3: Clustering Methods for Microarray Data, **MTCC, Auditorium**

Chair/Org: Hans Bock, University of Aachen

Guenther Walther, R. Tibshirani, D. Botstein, and P. Brown, Stanford University, “On Cluster Validation by Prediction Length”

Heather Turner, Trevor Bailey, and Wojtek Krzanowski, University of Exeter, “The Plaid Model: Some Enhancements and Extensions”

Alexander Schliep, Christine Steinhoff, and Alexander Schoenhuth, Max Planck Institute, “Inference of Groups in Gene Expression”

C-1: Approaches to Contingency Tables, **Alumni Lounge, HUB**

Chair: Michel Tenenhaus, HEC Business School

Stefano De Cantis and A. M. Oliveri, University of Palermo, “An Overview of Collapsibility”

François Bavaud, University of Lausanne, “Generalized Factor Analysis for Contingency Tables”

V. Loonis, Ecole Nationale de Statistique et de l’Administration Economique, “Simultaneous Row and Column Partitioning in Several Contingency Tables”

Ana Lorga da Silva, Universidade Lusofona de Humanidades e Tecnologias, Gilbert Saporta, CNAM, and Helena Bacelar-Nicolau, University of Lisbon, “Missing Data and Imputation Methods in Partition of Variables”

E. Acuña and Caroline Rodriguez, University of Puerto Rico, “The Treatment of Missing Values and Its Effect on Classifier Accuracy”

Mark de Rooij, Leiden University, “Scaling the Odds”

C-2: Symbolic Data Analysis I, **Hermann Lounge, HUB**

Chair: André Hardy, University of Namur,

Barbara Japelj Pavešič, Educational Research Institute, Simona Korenjak-Černe, University of Ljubljana, and V. Batagelj, University of Ljubljana, “An Application of Symbolic Clustering Methods on the TIMSS Dataset”

Jean-Paul Rasson and Pascale Lallemand, Facultés Universitaires Notre-Dame de la Paix, “Clustering Trees Applied to Symbolic Objects”

O. Rodríguez, University of Costa Rica, W. Castillo, University of Costa Rica, E. Diday, University of Paris IX–Dauphine, and J. González, University of Costa Rica, “Correspondence Factorial Analysis for Symbolic Multi-Valued Variables”

Marc Csernel, INRIA, Francisco de A. T. Carvalho, Universidade Federal de Pernambuco, and Yves Lechevallier, INRIA-Rocquencourt, “Partitioning of Constrained Symbolic Data”

Ole Nordhoff, Institut für Statistik und Wirtschaftsmathematik, “Expectation of Random Sets and the ‘Mean Values’ of Interval Data”

Sérgio R. de M. Queiroz and F. A. T. de Carvalho, Universidade Federal de Pernambuco, “A Symbolic Model-Based Approach for Making Collaborative Group Recommendations”

**12:30-1:15** Transversal Meeting on Symbolic Data Analysis, **Hermann Lounge, HUB**

Chair/Org.: Yves Lechevallier, INRIA-Rocquencourt

Edwin Diday, University of Paris–Dauphine, “Introduction to Symbolic Data Analysis”

Paula Brito, University of Porto, “The ASSO Project”

**2:00-3:00** Special Invited Talks

Chair: Lee Wilkinson, SPSS, Inc.

Pedro Domingos, University of Washington, “A General Framework for Mining Massive Data Streams” **Ballroom, HUB**

Chair: Stan Wasserman, Indiana University

Anuska Ferligoj, University of Ljubljana, Vladimir Batagelj, University of Ljubljana, and Patrick Doreian, University of Pittsburgh, “Blockmodeling as a Clustering Problem” **Alumni Lounge, HUB**

Chair: Jean-Paul Rasson, University of Namur

Fionn Murtagh, Queen's University Belfast, "Thinking Ultrametrically"  
**MTCC, Auditorium**

**3:30-5:00 p.m** Sessions (Short Session)

**I-4: Latent Variables, MTCC, Auditorium**

Chair: C. Anderson, University of Illinois at Urbana-Champaign

Jeremy Nadolski and Kert Viele, University of Kentucky. "The Role of Latent Variables in Model Selection Accuracy"

Patrick J. F. Groenen, Erasmus University, and Michael W. Trosset, College of William & Mary, "Fast Multidimensional Scaling of Large Data Sets"

**C-3: New Approaches in Cluster Analysis, Alumni Lounge, HUB**

Chair: Maurizio Vichi, University of Rome

Halima Bensmail and Hamparsum Bozdogan, University of Tennessee at Knoxville, "Bayesian Unsupervised Clustering for Mixed Data with Missing Observations"

J.-P. Rasson and Laurent Sancinito Spito, University of Namur, "Partitioning Non-Convex Clusters Using Alpha-Shapes"

Javier Trejos, Alex Murillo, and Eduardo Piza, University of Costa Rica, "Clustering by Ant Colony Optimization"

Jean-Yves Pirçon and Jean-Paul Rasson, University of Namur, "The Last Step of a New Divisive Monothetic Clustering Method: The Gluing-Back Criterion"

**C-4: New Kinds of Problems, Hermann Lounge, HUB**

Chair: Jean-Pierre Barthélemy, GET-ENST Bretagne

Russell Steele, McGill University, and Luis Garcia, Virginia Tech University, "Algebraic Geometry and Model Selection"

Matevz Bren, University of Maribor, and Vladimir Batagelj, University of Ljubljana, "Compositional Data Analysis with R"

Bernard Fichet, University of Aix Marseille II, "The Cartesian Product of  $\lambda$ -Quasi-Hierarchies"

Seungmin Nam, Kiwoong Kim, Sinsup Cho, Seoul National University, and Inkwon Yeo, Chonbuk National University, "A Bayesian Analysis Based on Beta-Mixtures for Software Reliability Models"

# Friday, July 16

## 9:00-10:00 a.m. Plenary **Ballroom, HUB**

Chair: Doug Carroll, Rutgers University

Gilles Celeux, INRIA, “Choosing a Model for Purposes of Classification”

## 10:30-12:30 Sessions

### I-5: Multiway Data Analysis, **Ballroom, MTCC**

Chair/Org.: Henk Kiers, University of Groningen

Iven van Mechelen, Catholic University of Leuven, “Three-Way Classification Models: An Overview”

Age Smilde, University of Amsterdam and TNO Nutrition and Food Research, Jeroen Jansen, Huub C. J. Hoefsloot, Jan van der Greef, and Marieke E. Timmerman, “Multiset Methods for Longitudinal Metabolomics Data”

Maurizio Vichi and Roberto Rocci, University “La Sapienza” and University “Tor Vergata” of Rome, “Multimode Clustering”

### I-6: Classification and Clustering in Astronomy, **Ballroom, HUB**

Chair/Org.: Bill Shannon, Washington University at St. Louis

Vicki Laidler, Computer Sciences Corporation, Space Telescope Science Institute, “A Tale of Two Probabilities: Assessing Accuracy on Classified Astronomical Data”

Kiri Wagstaff, JPL, “Clustering with Missing Values: No Imputation Required”

Markus Demleitner, University of Heidelberg, Michael Kurtz, Harvard University, et al. “Automated Resolution of Noisy Bibliographic References”

Disc.: Fionn Murtagh, Queen’s University Belfast

### I-7: Market Research, Data Mining, and Web Mining Interface **Alumni Lounge, HUB**

Chair/Org: Wolfgang Gaul, University of Karlsruhe

Daniel Baier, University of Cottbus “Pharmaceutical eDetailing and Market Sementation in the Age of the Internet”

Reinhold Decker, University of Bielefeld “Self-Controlled Growing Neural Networks and Their Application to Pattern Representation in Data Analysis”

Born, Stefan, Liebig University, and Lars Schmidt-Thieme, University of Freiburg “Optimal Discretization of Quantitative Attributes for Association Rules”

C-5: Graphs and Consensus, **Auditorium, MTCC**

Chair: Javier Trejos, University of Costa Rica

Helena Brás Silva, Paula Brito, and Joaquim Costa, University of Porto, “Clustering and Validation Using Graph Theory: Application to Data Sets with Numerical and Categorical Values”

Alain Guénoche, Institute de Mathématiques de Luminy–CNRS, “Clustering By Vertex Density in a Graph”

Vladimir Batagelj and Matjaž Zaveršnik, University of Ljubljana, “Islands, an Approach to Clustering in Large Networks”

Christophe Osswald and François Brucker, GET-ENST Bretagne, “Circular Dissimilarities and Hypercycles”

Guy Cucumel, Université du Québec à Montréal, “Average Consensus and  $L_\infty$ -consensus: Comparison of Two Consensus Methods for Ultrametric Trees”

Robert C. Powers, University of Louisville, “Consensus Functions on Hierarchies that Satisfy Independence and Pareto”

C-6: New Ideas in Classification, **Hermann Lounge, HUB**

Chair: Myung-Hoe Huh, Korea University

Seong Keon Lee, Chuo University, “On Classification and Regression Trees with Multiple Responses”

Bart Jan van Os, Leiden University, “Globally Optimal Partitioning by Dynamic Programming”

Jan Schepers and Iven Van Mechelen, University of Leuven, “Three-Mode Partitioning: An Evaluation of Three Algorithms”

Rafik Abdesselam, Université de Caen Basse-Normandie, “Sequential Multiple Discriminant Analysis”

Koji Kurihara, Okayama University, “Classification of Geospatial Lattice Data and Their Graphical Representation”

Kaddour Bachar, Ecole Supérieure des Sciences Commerciales d’Angers, and Israël Lerman, University of Rennes 1, “Fixing Parameters in the Constrained Hierarchical Classification Method: Application to Digital Image Segmentation”

**2:00-3:00** Special Invited Talks

Chair: Balaji Krishnapuram, Duke University

Vladimir Vapnik, AT&T, Title TBA, **Auditorium, MTCC**

Chair: Yves Lechevallier, INRIA-Rocquencourt

Lynne Billard, University of Georgia, “Symbolic Data Analysis: An Overview of Principles and Some Descriptive Statistics” **Ballroom, HUB**

Chair: William H. E. Day, Port Maitland

Bruno Leclerc, EHESS, “Consensus of Classification Systems, with Adams’ Results Revisited” **Alumni Lounge, HUB**

**3:30-5:30 p.m** Sessions

I-8: Roundtable for IFCS Education Committee: “Teaching and Training Classification and Data Analysis in a Multidisciplinary and Interdisciplinary Context” **Hermann Lounge, HUB**

Org.: Helena Bacelar-Nicolau, University of Lisbon

Chair: Henk Kiers, University of Groningen

Henk Kiers, University of Groningen, “Teaching Data Analysis to First Years: How to Avoid that They Only Learn the *How* and Not the *Why*”

Fionn Murtagh, Queen’s University Belfast, “New Research Results and Best Practice in Classification and Data Analysis in Established Disciplines: The Case of Astronomy and Astrophysics”

Krzysztof Jajuga, Wroclaw University of Economics, “Teaching Statistics for Economics and Finance Majors and Practitioners”

Gilbert Saporta, CNAM Paris, “A Brief Report on Teaching Statistics with the Internet at an Introductory Level”

Peter Bryant, University of Colorado at Denver, “How Would You Go About Teaching Principal Components? Developing a *small case*”

Iven Van Mechelen, Catholic University of Leuven, “Three Major Obstacles/Challenges for Teaching Classification Models and Methods”



I-9: Data Mining I, **Ballroom, HUB**

Org.: David Banks, Duke University

Chair: Greg Constantine, University of Pittsburgh

Kath Bogie, Jiayang Sun, and Xiaofeng Wang, Case Western University, “Mining Huge-p-Small-n Data and New Image Registration Procedures”

Balaji Krishnapuram, Duke University, “Autonomous Learning of Multi-Sensor Classifiers”

Ernest Fokoue, Ohio State University, “Variational Approximation for Gaussian Process Classifiers”

I-10: Statistical Methods in Immunology and Medicine, **Alumni Lounge, HUB**

Org: Greg Constantine, University of Pittsburgh

Chair: Taerim Lee, Korea National Open University

Gilles Clermont, Carson Chow, Greg Constantine, et al., University of Pittsburgh, “Mathematical Modeling of Acute Inflammation”

Gregory Constantine, University of Pittsburgh, “A Multidimensional Parameter Estimation Algorithm Based on Optimal Linear Codes”

Jennifer Pittman, Duke University, “Bayesian Analysis of Binary Trees and Model Averaging”

C-7: Text Analysis, **Ballroom, MTCC**

Chair: Stan Sclove, University of Illinois at Chicago

David I Holmes and Daniel W Crofts, The College of New Jersey, “The Diary of a Public Man: A Case Study in Traditional and Non-Traditional Authorship Attribution”

D. Birnbaum, University of Pittsburgh, and David Dubin, University of Illinois at Urbana-Champaign, “Measuring Similarity in the Contents of Medieval Miscellany Manuscripts”

Bei Yu, University of Illinois at Urbana-Champaign, “Writing Style Discrimination Between Native and Non-native English Writers”

Takahiro Saito, Hosei University, Kenji Nagasaka, Hosei University, and Tomokazu Takahashi, Aoyamagakuin University, “Classification Problems for Story Animations”

Mónica Bécue, Universitat Politècnica de Catalunya, J. Pagès, Institut Nacionales Supérieur, and Campo-Elías Pardo, Universidad Nacional de Colombia, “Analysis of Cross-Language Open-Ended Questions Through MFACT”

Georges Hébrail, ENST, and Yves Lechevallier, INRIA, “Building Small Scale Models of Multi-Entity Databases by Clustering”

C-8: Neural Nets and Variable Selection, **Auditorium, MTCC**

Chair: Keiji Yajima, Science University of Tokyo

Herbert K. H. Lee, University of California at Santa Cruz, “Priors for Neural Networks”

Reinhold Decker, University of Bielefeld, “Self-Controlled Growing in Neural Networks”

Tepei Shimamura, Hiroyuki Minami, and Masahiro Mizuta, Hokkaido University, “Nonlinear Regression Modeling Strategy Based on Radial Basis Function Networks and Absolute Shrinkage”

José Luis Espinoza, Instituto Tecnológico de Costa Rica, and Javier Trejos, CIMPA, “Genetic Variable Selection in Linear Regression”

Brieuc Connan-Guez, INRIA-Rocquencourt, and F. Rossi, University of Paris Dauphine, “Phoneme Discrimination with Functional Multi-Layer Perceptrons”

Raphaël Couturier, LIFC-IUT Belfort-Montbéliard, Régis Gras, and Fabrice Guillet, École Polytechnique de l’Université de Nantes, “Reducing the Number of Variables Using Implicative Analysis”

# Saturday, July 17

**9:00-10:00 a.m. Plenary, Ballroom, HUB**

Chair: F. R. McMorris, Illinois Institute of Technology

Mike Steel, University of Canterbury, New Zealand, “Phylogenetic Closure Operations and Homoplasy-Free Evolution”

**10:30-12:30 Sessions**

**I-11: Information Retrieval I., Ballroom, MTCC**

Org.: Nazli Goharian, Illinois Institute of Technology

Chair: Wai Gen Yee, Illinois Institute of Technology

Simona Balbi and Emilo Di Meglio, University of Naples, “Contributions of Textual Data Analysis to Text Retrieval”

W. Wu, University of Illinois at Urbana-Champaign, Clement Yu, University of Illinois at Chicago, and Weiyi Meng, State University of New York at Binghamton, “Database Selection for Longer Queries”

Ahu Sieg, Bamshad Mobasher, and Robin Burke, DePaul University, “Inferring User’s Information Context from User Profiles and Concept Hierarchies”

**I-12: Functional Magnetic Resonance Imaging, Auditorium, MTCC**

Org.: Mark Vangel, Massachusetts General Hospital and Harvard University

Chair: Israël Lerman, University of Rennes 1

Nicole Lazar, Carnegie Mellon University, “Are All fMRI Subjects Created Equal?”

Tom Nichols, University of Michigan, “Detecting a Conjunction of Alternatives: Finding Consistent Activations in Functional Neuroimaging Data using FDR”

Mark Vangel, Massachusetts General Hospital and Harvard University, “Combining Functional MRI Data on Multiple Subjects”

**I-13: Statistical Issues in Counterterrorism, Ballroom, HUB**

Chair/Org.: David Banks, Duke University

David Marchette, Naval Surface Warfare Center, “Iterative Denoising for Cross-Corpus Discovery”

Karen Kafadar, University of Colorado at Denver, and Cliff Spiegelman, Texas A&M University, “Forensic Analysis of Bullet Data”  
Michael D. Larsen, Iowa State University, “Issues in Record Linkage”

C-9: Applications in Business, **Armour Dining Room, HUB**

Chair: Andrzej Sokolowski, Cracow University of Economics

Chih-Chou Chiu, National Taipei University of Technology, and T.-S. Lee, Fu-Jen Catholic University, “Integrating Classification and Regression Trees with Multivariate Adaptive Regression Splines in Mining Customer Credit Data”

Sun-Soon Lee, Seoul National University, Hong-Seok Lee, Sung Kyun Kwan University, Joong-Hwan Lee, Needs I Co. Ltd., and Sung-Soo Kim, Korea National Open University, “Customer Segmentation using gCRM”

Carlos Lourenço and M. Cardoso, ISCTE, “Market Segmentation: A Comparison Between Mixture Models and Data Mining Techniques”

Katarzyna Kuziak, Wroclaw University of Economics, “Evaluation of Risk in the Index Option Pricing Model”

A. Chaturvedi, Capital One Services, J. Douglas Carroll, Rutgers University, and V. Cartwright, Capital One Services, “Three-Way, Three-Mode Hybrid Factor-Components Analysis Model”

Basavanneppa Tallur, University of Rennes 1, “The Hierarchical Clustering of Binary Attribute Data Based on Correlation with Application to Bioinformatics”

C-10: Issues in k-Means Clustering, **Alumni Lounge HUB**

Chair: Yoshiharu Sato, Hokkaido University

C. Cifaelli and L. Nieddu, University of Rome, “An Iterative k-Means-Like Algorithm for Statistical Face Recognition”

Arta Doci and Peter Bryant, University of Colorado at Denver, “An MDL Criterion for k-Means Clustering”

M.-H. Huh, Korea University, and Yong-Goo Lee, Chung Ang University, “Reproducibility Evaluation of k-Means Clustering”

Douglas Steinley, University of Illinois at Urbana-Champaign, “Standardizing Variables in k-Means Clustering”

Aïcha El Golli, INRIA-Rocquencourt, Brieuc Conan-Guez, and F. Rossi, University of Paris–Dauphine, “A Self-Organizing Map for Dissimilarity Data”

Myung-Hoe Huh, Korea University, “Enhancing Self-Organizing Maps for Statistical Use”

**2:00-3:00** Special Invited Talks

Chair: Willem Heiser, Leiden University

Carey Priebe, Johns Hopkins University, “A Hierarchical Methodology for Classification Problems with Skewed Priors” **Ballroom, HUB**

Chair: Jennifer Pittman, Duke University

Taerim Lee, Korea National Open University, “A Tree-Structured Survival Model for AIDS in Korea” **Auditorium, MTCC**

Chair: David Dubin, University of Illinois at Urbana-Champaign

Regina Liu, Rutgers University, “Mining Massive Text Data and Developing Tracking Statistics” **Alumni Lounge, HUB**

**3:30-5:30 p.m** Sessions

I-14: Information Retrieval II, **Ballroom, HUB**

Org.: Nazli Goharian, Illinois Institute of Technology

Chair: Wai Gen Yee, Illinois Institute of Technology

Maojin Jiang, Eric Jensen, Steve Beitzel, and Shlomo Argamon, Illinois Institute of Technology, “Choosing the Right Bigrams for Information Retrieval”

Tao Tao and ChengXiang Zhai, University of Illinois at Urbana-Champaign, “A Mixture Clustering Model for Pseudo Feedback in Information Retrieval”

E. Graubins and David Grossman, Illinois Institute of Technology, “Applying Hybrid Modeling to Predict the Stock Market”

I-15: Dimension Reduction, **Auditorium, MTCC**

Chair: Masahiro Mizuta, Hokkaido University

Jeff Solka, Naval Surface Warfare Center, “Modern Geometric Methods for Dimensionality Reduction”

Xiaodong Lin, Statistics and Applied Mathematics Institute, and Yu Zhu, Purdue University, “Degenerate Expectation-Maximization Algorithm for Local Dimension Reduction”

Ulas Akkucuk and Doug Carroll, Rutgers University, “Mapping of Nonlinear Manifolds: A Study of ISOMAP and a Version of PARAMAP”

I-16: Issues in Statistical Metrology, **Alumni Lounge, HUB**

Chair: Patrick Groenen, Erasmus University of Rotterdam

Kevin Coakley, National Institute of Standards and Technology, “A Classification Problem in Neutrino Physics”

Tony Kearsley and Luis Melara, National Institute of Standards and Technology, “Nonlinear Programming and Multi-Dimensional Scaling, I”

Tony Kearsley and Luis Melara, National Institute of Standards and Technology, “Nonlinear Programming and Multi-Dimensional Scaling, II”

C-11: Applications in Medicine, **Armour Dining Room, HUB**

Chair: Alain Guénoche, University of Marseilles

F. Giannandrea, Catholic University of Sacred Heart, and D. F. Iezzi, University “La Sapienza”, “A Method to Classify Hospital Workers with Lower Back Pain”

F. Goupil-Testu, M. Touati, I. Wagne, Université of Paris - IX Dauphine, A. Pelc, CCMSA, and E. Diday, Université of Paris - IX Dauphine, “Analysis of Drug Dispensing with Symbolic Methods”

Hodaka Numasaki, H. Harauchi, Y. Ohno, Osaka University, K. Inamura, Kansai University of International Studies, S. Kasahara, M. Monden, and M. Sakon, Osaka University, “Application of Spectrum Analysis and Sequence Relational Analysis for the Medical Staff’s Job Classification: Use of Time Factor and Behavior Factors for Job Workflow”

Vahan Grigoryan, Donald Chiarulli, and Milos Hauskrecht, University of Pittsburgh, “Subject Filtering for Passive Biometric Monitoring”

Nicholas I. Bohnen, Marius Buliga, and Gregory M. Constantine, University of Pittsburgh, “Classification of the State of Parkinsonism by Using Electronic Force Platform Measures of Balance”

Vladimir Makarenkov, Alix Boc, and Abdoulaye Baniré Diallo, University of Quebec, “Determining Horizontal Gene Transfers in Species Classification: Unique Scenario”

C-12: New Ideas in Cluster Analysis, **Ballroom, MTCC**

Chair: Yong-Goo Lee, Chung Ang University

S.-S. Chae, Daejeon University, C. Kim, and W. D. Warde, Oklahoma State University, “Cluster Analysis Using Robust Estimators”

Patrice Bertrand, GET-ENST Bretagne, and Ghazi Bel Mufti, ESSEC, “Loevinger’s Measures of Rule Quantity for Assessing Cluster Stability”

María Teresa Gallegos and Gunter Ritter, Universität Passau, “A Break-point Analysis for Clustering”

Gunter Ritter and María Teresa Gallegos, Universität Passau, “Clustering of Contaminated and Ambiguous Data”

Sarnath Ramnath, M. H. Khan, and S. Shams, St. Cloud State University, “New Approaches for Sum-of-Distances Clustering”

Akinobu Takeuchi, Jissen Women’s University, T. Saito, Tokyo Institute of Technology, and H. Yadohisa, Kagoshima University, “Extension of Average Linkage Clustering to Asymmetric Data”

# Sunday, July 18

**9:00-10:00 a.m. Plenary, Ballroom, HUB**

Chair: Phipps Arabie, Rutgers University

David Hand, Imperial College, “Academic Obsessions and Classification Realities: Ignoring Practicalities in Supervised Classification”

**10:30-12:30 Sessions**

**I-17: Data Mining II, Ballroom, MTCC**

Chair/Org.: Ernest Fokoue, Ohio State University

M. Markatou, H. Tian, S. Biswas, and G. Hripsack, Columbia University, “An Analysis of the Influence of Correlation on the Cross-Validation Estimates of Performance for Learning Algorithms”

Leanna House and David Banks, Duke University, “Cherry-Picking as a Robustness Tool”

Helen Zhang, North Carolina State University, “Unified Multiclass Proximal Support Vector Machines”

**I-18: Data Analysis in Financial Risk Management, Ballroom, HUB**

Org.: Krzysztof Jajuga, Wroclaw University of Economics

Chair: Anil Chaturvedi, Capital One Services

Krzysztof Jajuga, “Data Analysis and Financial Risk Management: Overview of Modern Approaches”

Michael Ong, Illinois Institute of Technology, “The Use and Ab-Use of Financial Data”

H. Locarek-Junge, Dresden University of Technology, “Estimation of Tail Coefficients and Extreme Correlations for Market and Credit Risk: Problems, Pitfalls, and Possible Solutions”

**C-13: New Ideas for Binary Data, Alumni Lounge, HUB**

Chair: Yongil Kim, Chung Ang University

Jean-Pierre Barthélemy, GET-ENST Bretagne, “Binary Clustering and Boolean Dissimilarities”

François Brucker, GET-ENST Bretagne, “On a Bijection Between Dissimilarities and Binary Realizations”

Yoshiharu Sato, Hokkaido University, “Clustering Binary Data Using a Transformation into Directional Data”



Ingo Ruczinski, Johns Hopkins University, “Improvements for Logic Regression”

S. Gardner and N. le Roux, University of Stellenbosch, “Modified Biplots for Enhancing Two-Class Discriminant Analysis”

C-14: Symbolic Data Analysis II, **Hermann Lounge, HUB**

Chair: Paula Brito, University of Porto

F. Afonso, University of Paris–Dauphine, Lynne Billard, University of Georgia, and Edwin Diday, University of Paris–Dauphine, “Symbolic Linear Regression with Taxonomies”

André Hardy and P. Lallemand, University of Namur, “Clustering of Symbolic Objects Described by Multi-Valued and Modal Variables”

Marie Chavent, University of Bordeaux, “A Hausdorff Distance Between Hyper-Rectangles for Clustering Interval Data”

Chérif Mballo, ESIEA, and Edwin Diday, University of Paris–Dauphine, “Kolmogorov-Smirnov for Decision Trees on Interval and Histogram Variables”

Mohamed Mehdi Limam, University of Paris–Dauphine, Edwin Diday, University of Paris–Dauphine, and Suzanne Winsberg, Institut de Recherche et Coordination Acoustique et Musique, “Probabilistic Allocation of Aggregated Statistical Units in Classification Trees for Symbolic Class Description”

Renata M. C. R. de Souza, Cidade Universitaria–Brazil, F. A. T. de Carvalho, Universidade Federal de Pernambuco, C. P. Tenório, and Y. Lechevallier, INRIA-Rocquencourt, “Dynamic Cluster Methods for Interval Data Based on Mahalanobis Distances”

C-15: Latent Variables and Related Topics, **Auditorium, MTCC**

Chair: Akinori Okada, Rikkyo University

José G. Dias, Instituto Superior de Ciências do Trabalho e da Empresa, “Controlling the Level of Separation of Components in Monte Carlo Studies of Latent Class Models”

G. Luo and D. Andrich, Murdoch University, “Weighted Likelihood Estimation of Person Locations in an Unfolding Model for Polytomous Responses”

Kohei Adachi, Ritsumeikan University, “An Oblique Procrustes Method for Simultaneously Rotating Component and Loading Matrices”

Mohamed Nadif, Université de Metz, and Gérard Govaert, Université de Technologie de Compiègne, “Another Version of the Block EM Algorithm”

T. Murakami, Nagoya University, “Direct Procrustes Method and Perfect Congruence”

**2:00-4:00 p.m** Sessions

**I-19: Data Mining III, Ballroom, MTCC**

Chair/Org.: Xiaodong Lin, Statistics and Applied Mathematical Science Institute

Stan Young, National Institute of Statistical Sciences, “Using Metabolomic Data to Predict Disease Status”

Paul Marriott, University of Waterloo, “Complexity Issues in the Geometry of Parametric Models”

J. Hughes-Oliver, North Carolina State University, “Twins and High Dimensional Data: Can Leave-one-out Cross Validation Survive?”

**I-20: Memorial Session for Chikio Hayashi, Auditorium, MTCC**

Org.: Noboru Ohsumi, Institute of Statistical Mathematics

Chair: Alfredo Rizzi, University of Rome

Noboru Ohsumi, Institute of Statistical Mathematics, “Memories of Chikio Hayashi and His Great Achievement”

Ludovic Lebart, ENST, “Validation Technique in Correspondence Analysis”

Hans-Hermann Bock, “Classification in the Life Span of Chikio Hayashi”

Yasumasa Baba and Noboru Ohsumi, Institute of Statistical Mathematics, “Chikio Hayashi and Data Science in Japan”

**C-16: Odds and Ends, Alumni Lounge, HUB**

Chair: Bernie Harris, University of Nebraska-Lincoln

Francisco de A. T. Carvalho, Universidade Federal de Pernambuco, Yves Lechevallier, INRIA-Rocquencourt, Renata M. C. R. de Souza, Cidade Universitaria–Brazil, “A Dynamic Clustering Algorithm Based on  $L_r$ -Distances for Quantitative Data”

Kutluhan Kemal Pak and Edwin Diday, Univeristé Paris–IX Dauphine, “Construction of Spatial Hierarchies”

Margaret Polinkovsky, Duke University, “Analysis of Key Comparisons”

Anil Chaturvedi, Capital One Services, J. Douglas Carroll, Rutgers University, and Vicki Cutwright, Capital One Services, “HIDCLAN: A General Approach to Finding Hidden Clusters and Latent Classes in Multivariate Data Analysis”

**4:30-5:30** Special Invited Talks

Chair: Herbert K. H. Lee, University California at Santa Cruz

Masahiro Mizuta, Hokkaido University, “Relative Projection Pursuit and Its Applications” **Ballroom, HUB**

Chair: Otto Opitz, University of Augsburg

Wolfgang Gaul, University of Karlsruhe, “Web Mining and Its Interrelations to Classification Data Mining and Market Research” **Auditorium, MTCC**

Chair: Lynne Billard, University of Georgia

Edwin Diday, University of Paris–Dauphine, “Introduction to Spatial Classification” **Alumni Lounge, HUB**

**6:00-8:00** Closing Session, **Ballroom, HUB**

Chair: Vladimir Batagelj, University of Ljubljana

David Wishart, University of St. Andrews, “Classification of Single Malt Whiskies by *Flavour*”